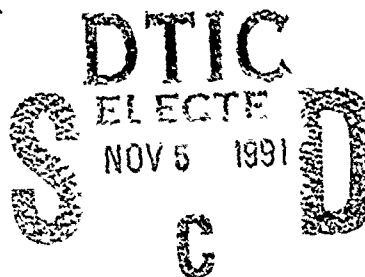


AD-A242 781



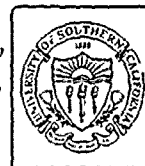
Eve M. Schooler  
Stephen L. Casner  
Jon Postel



ISI Reprint Series  
ISI/RS-91-286  
August 1991

2

University  
of Southern  
California



Multimedia Conferencing:  
Has it come of age?

Reprinted from the  
*Proceedings 24th Hawaii International  
Conference on System Sciences*,  
Vol.3, pp.707-716 (January 1991).

91 1104 154

DISCLOSURE STATEMENT A

Approved for public release;  
Distribution Unlimited

91-15075



INFORMATION  
SCIENCES  
INSTITUTE



213/822-1511

4676 Admiralty Way/Marina del Rey/California 90292-6695

## REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S) ISI/RS-91-286			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION USC/Information Sciences Institute		6b. OFFICE SYMBOL (if applicable)	7a. NAME OF MONITORING ORGANIZATION Department of the Army		
6c. ADDRESS (City, State, and ZIP Code) 4676 Admiralty Way, Marina del Rey, CA 90292-6695			7b. ADDRESS (City, State, and ZIP Code) Directorate of Contracting, ATTN: ATZS-DKM Post Office Box 748, Fort Huachuca, AZ 85613-0748		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION DARPA		8b. OFFICE SYMBOL (if applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DARPA (Ft. Huachuca); DABT63-91-C-0001		
8c. ADDRESS (City, State, and ZIP Code) 3701 N. Fairfax Drive Arlington, VA 22203-1714			10. SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.
					WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) Multimedia Conferencing: Has it come of age? (Unclassified)					
12. PERSONAL AUTHOR(S) Schooler, E.M.; Casner S.; Postel, J.					
13a. TYPE OF REPORT Research Report		13b. TIME COVERED FROM TO		14. DATE OF REPORT (Year, Month, Day) 1991, August	
15. PAGE COUNT					
16. SUPPLEMENTARY NOTATION Reprinted from the <i>Proceedings 24th Hawaii International Conference on System Sciences</i> , Vol.3, pp. 707-716 (January 1991).					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
09	02		Multimedia, teleconferencing, collaborative work, groupware		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)					
<p>Multimedia conferencing continues to be an active research area with a diversity of systems being developed for a variety of situations. In particular, multimedia conferencing is central to the idea of the Collaboratory, an electronic environment for conducting science. As a vehicle for telecollaboration, multimedia conferencing promises to provide a meeting place for those needing to work cooperatively from afar. This paper reviews current research projects in telecollaboration. We first frame the discussion of multimedia conferencing with a nomenclature and taxonomy. We compare local versus remote conferencing, touching on issues in system architecture and network communications requirements. We then examine the recurring problems researchers have observed and the solutions they have chosen. Finally, we speculate that, with the advent of new technologies and with a growing sensitivity to human factors, if multimedia conferencing has not already come of age, it now has the opportunity.</p>					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION Unclassified		
22a. NAME OF RESPONSIBLE INDIVIDUAL Kathleen McLaughlin/Celeste Anderson			22b. TELEPHONE (Include Area Code) 213-822-1511		22c. OFFICE SYMBOL

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

Eve M. Schooler  
Stephen L. Casner  
Jon Postel

University  
of Southern  
California



## Multimedia Conferencing: Has it come of age?

Reprinted from the  
*Proceedings 24th Hawaii International  
Conference on System Sciences*,  
Vol.3, pp.707-716 (January 1991).

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



INFORMATION  
SCIENCES  
INSTITUTE



213/822-1511

4676 Admiralty Way/Marina del Rey/California 90292-6695

# Multimedia Conferencing: Has it come of age?

Eve M. Schooler, Stephen L. Casner, and Jon Postel

USC/Information Sciences Institute  
4676 Admiralty Way  
Marina del Rey, CA 90292

## 1 Introduction

Multimedia conferencing has come of age. Or, has it? Proponents of teleconferencing have made statements like this since as early as the 1920's when the idea of video conferencing debuted [20], and echoed this pronouncement again during the 1960's when AT&T introduced its PicturePhone [3] at the 1964 World's Fair. Marketing forecasts of the 1970's promised the teleconferencing revolution [35] and touted "videoconferencing as a revolutionary concept on the brink of success" [9]. One must wonder, are these assertions any more true now than they were then?

Teleconferencing is hardly a novel concept. Yet, it has consistently fallen short of expectations as an effective means of communication. Grudin [17] attributes this to the technologically-driven nature of the pursuit and paraphrases a colleague who sees this shortcoming as "technology searching for a need". Egido's articulate discussion of its failures points to factors lying beyond the scope of technology, such as psychological and sociological ones, and argues that the casting of electronic communication in the image of face-to-face meetings has stood in the way of developing multimedia conferencing technology to its fullest potential [9]. Bikson lobbies for systems more attuned to group processes, taking the stance that system builders must consider the tools and technology already in place, as well as individual preferences [4].

Despite valid skepticism, multimedia conferencing continues to be an active research area with a diversity of systems being developed for a variety of situations. In particular, multimedia conferencing is central to the idea of the Collaboratory, an electronic environment for conducting science. As a vehicle for telecollaboration, multimedia conferencing promises to provide a meeting place for those needing to work cooperatively from afar. This paper reviews current research projects in telecollaboration and how they have addressed the aforementioned criticisms. We first frame the discussion of multimedia conferencing with a nomenclature and taxonomy. We compare local versus remote conferencing, touching on issues in system architecture and network communication requirements. We then examine the recurring problems researchers have observed and the solutions they have chosen. Finally we speculate that, with the advent of new technologies and with a growing sensitivity to human factors, if multimedia conferencing has not already come of age, it now has the opportunity.

## 2 Nomenclature and Taxonomy

A wide range of work falls under the heading *multimedia conferencing*. In the broadest sense it is the use of mixed media for group collaboration. The term *multimedia* itself has a wide range of meanings. We use it to refer to a collection of computer-based media such as text, structured graphics, bitmaps, facsimile formats and spreadsheets, plus real-time voice and video. Some multimedia documents also embed audio segments and video stills or animations.

Several variables help to differentiate conferencing systems; figure 1 lists pairs of contrasting characteristics. One may think of these variables as forming a multidimensional space with each system falling somewhere in that space (admittedly, some points within the space are uninteresting).

Perhaps the most basic division is between *synchronous* and *asynchronous* conferencing. While both forms of conferencing cater to multiple users, synchronous conferencing is intended for simultaneous users who have real-time interactions, while asynchronous conferencing systems, such as multimedia electronic mail [29] and structured

messaging systems [22], provide non-real-time communication. In this paper, we concentrate our discussion on synchronous conferencing systems.

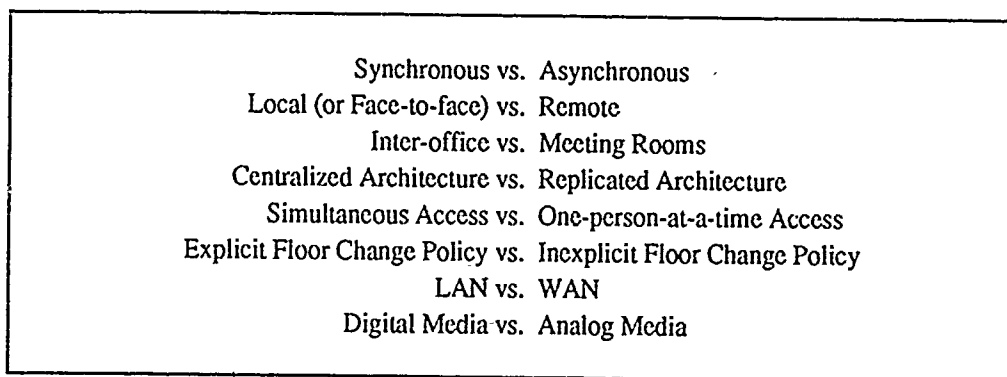


Figure 1. Conferencing Characteristics

Another fundamental distinction is *local* face-to-face computer-augmented meetings [26, 36] versus *remote* meetings for which a real-time voice and/or video channel is required [7, 23, 31, 33]. These live media may be carried in *digital* [33] or *analog* [31] form. Some remote conferencing systems are designed for *inter-office* collaboration [31] while others are for conferences between special *meeting rooms* [33]. Some remote systems will only operate with the low delays seen across a local area network (LAN) [26, 31, 36], while others can tolerate the longer delays of a more geographically dispersed wide area network (WAN) [7, 23, 33].

Most multimedia conferencing systems include computer-based tools to support group collaboration, or *groupware* [7, 11, 12, 23]. Some groupware applications are considered meeting tools [36] that aid the meeting process itself, like voting or brainstorming, and others aim to bring subject matter into the meeting [7, 23]. Computer conferencing tools may be classified as having either a *centralized* [23] or *replicated* [7] architecture. The centralized approach is based on the execution of the application at one site with input forwarded from whichever site has the floor to the site where the application executes and all output broadcast to the other sites. By comparison, a fully replicated architecture runs a copy of the application at each site in the conference. Input from the site with the floor is broadcast to the other participating sites and output is generated locally at each site. Replicated systems minimize the input to output delay for the participant with the floor, but are harder to construct [7, 24].

Within groupware applications, a range of *floor policies* are used. Some systems allow simultaneous access to the shared workspace by multiple users [36], while others only allow one user to alter the work area at a time [7, 23]. To obtain the "floor", one may be required to take an explicit action, like the selection of a special function key [26]. Less restrictive systems allow any keyboard or mouse activity to signal a floor change [7]. More recently, some systems are providing a range of policies to fit the different types of meetings that arise [7, 32].

As conferencing systems become more sophisticated, they may support *venue-agility* [18], that is they may allow users to operate in multiple points of the multidimensional space. Such a system may support a move between synchronous modes (e.g., collaborative editing) and asynchronous modes (e.g., electronic mailings of these edits), or allow the transition from working stand-alone, to working with one other person, to working with a group of people.

### 3 Current Systems

To elaborate on the conferencing nomenclature and taxonomy, we discuss a sample of current multimedia conferencing systems which might contribute to the realization of the Collaboratory. We note the variation in capabilities, architecture and results gleaned from each. Many of the projects in the last decade have been inspired by the seminal ideas of Bush, whose hypothetical Memex [5] predates today's hypertext systems as multi-user repositories of information. To an even larger degree many were influenced by Engelbart, whose NLS/AUGMENT [12, 11] was one of the first systems to use computers for group collaboration. Engelbart's system not only operated asynchronously to support electronic mail, but provided synchronous modes as well, addressing the logistics of terminal linking, sharing of files and floor control.

### 3.1 The Colab

Xerox PARC's Colab [36] was designed as an experimental computer-equipped meeting room that allowed small groups (2-6 people) to focus on problem solving in face-to-face meetings. The room was configured with personal computers which were set around a U-shaped table and were connected over a LAN. The table faced a large, touch-sensitive screen at the front of the room. This setting gave rise to several computer-based meeting tools including Cognoter, a tool to collectively organize ideas for presentations, and Argnoter, an argument spreadsheet for proposals.

Motivations behind Colab were studies that suggest office workers spend between 30-70% of their time in meetings [27], that the presence of computers in meetings is minimal, and that software which runs on computers is typically geared toward individuals rather than groups. Also, certain tasks regularly found in meetings are well suited to computers, computers can crisply display, manipulate, store and redisplay information better than a blackboard can. The Colab developers concentrated on face-to-face conferencing since these kinds of collaborations most often occurred in their research group.

The Colab's replicated architecture employed a distributed database. Several different control models (e.g., centralized locks, token passing) were considered for synchronization of shared data. To minimize delays to users, a cooperative model was chosen. Each machine had a copy of the database and changes were installed by broadcasting each modification without any synchronization. Race-conditions were mitigated by the use of visual and verbal cues. The graphical user interface would gray out portions of the screen to provide a busy signal if it contained data being modified by another group member. The participants also relied on verbal negotiation with other group members before altering shared data. These techniques effectively supported the floor policy that all users had simultaneous access to the shared workspace.

A continued theme has been to explore how to visually display shared simultaneously-accessible workspaces. The initial approach was to enforce a strict What-You-See-Is-What-I-See (WYSIWIS) policy for the display. This was relaxed to allow personalized window layouts including private windows. It was also found that multiple cursors (one for each participant) on the screen led to too much confusion so pointers were made visible only on request. However, personalized views of a public window may also cause confusion if one participant points to data which does not appear in another participant's view; the ability to "snap back" to a consistent shared view is critical.

### 3.2 The Capture Lab

Another LAN-based face-to-face conferencing facility was the University of Michigan's Capture Lab. Built on Colab findings, it experimented with human factors issues like seating arrangements of the users, the field of view between meeting participants, and the protocols used for the exchange of information between individuals [26].

The specialized meeting room was designed to look like a conventional conference room, save for a computer workstation per person and a shared electronic blackboard at the front of the room. This meant the use of an oval conference table with inlaid computer monitors. They found that attendees participated more in these meetings than when they met at a U shaped table and attributed this phenomenon to better eye contact between participants. Having witnessed the effect of a room configuration change on the interplay of the meeting participants, they addressed other room layout concerns, like providing equal visibility of the front screen from anywhere at the table, or what they refer to as *seating equality*, the layout of the personal computers so as not to occlude others; and room and table coloration to overcome the bulkiness of the workstations to make them appear less obtrusive.

In the Colab project, since all participants could enter data at once, they became totally absorbed in their typing, so eye contact and verbal exchanges diminished. To counteract this problem, Capture Lab made a departure from Colab style WYSIWIS. The individual workstations maintained private workspaces for each user with the front screen designated as the global working area. Pre-emptive sequential control of the shared electronic blackboard was provided through the use of a function key on the keyboard, although participants were often found to verbally discuss floor passing beforehand as well. Users could add private data to the shared work area once they received the floor.

Several unexpected phenomena were observed. The electronic blackboard captured group attention so much in meetings that changes made to it often caused individuals who were speaking to lose their train of thought. Also,

three types of meetings were noted to take place depending on how comfortable everyone was with the software, on the relative typing skills of conferees, and on the degree of preparation and formality in the meeting. The Capture Lab creators expected *interactive meetings* with equal amounts of participation by all individuals, but found these happened rarely. The most common type was the *rotating scribe meeting* where those most dextrous acted as scribes to the shared electronic bulletin board. Lastly, they found that more formal meetings with a *designated scribe* generally broke down into rotating scribe meetings.

### 3.3 MMConf

BBN's MMConf system [7, 14] differs from the Colab and Capture Lab projects in that it was designed to accommodate remote conferencing -- distributed, real time group interactions of an inter-office or inter-meeting-room nature. As a conferencing umbrella program, MMConf supports a variety of applications in multi-user mode, ranging from a simple sketch tool, to a multimedia editor, a presentation tool, and video map and database browsers. These applications bring subject matter into meetings, rather than act as meeting-support tools. Because MMConf is used among remote locations, a voice channel is required to substitute for face-to-face speech. Conventional telephone conference calls have been used as well as a packet-switched video teleconferencing system (see the section *The DARPA Multimedia Conferencing Project*).

Applications running under MMConf use *gavel passing* for floor control, one participant has the floor at any given time. A floor request is implicit in any keyboard or mouse button input. The site with the floor hands off the floor when requested. Unlike Capture Lab, there is no explicit function key to control this and, unlike Colab, only one person has the floor at a time. The conferencing system's audio is often used to negotiate who should take the floor next since, without verbal agreement, a flurry of retries sometimes results. MMConf manages only certain windows on a user's workstation. The other windows are deemed private and allow users to work independently. Data from private windows can be imported into the global work area through use of conventional window cut-and-paste functions.

MMConf uses a replicated conferencing architecture, requiring each site to have its own copy of the requisite files, be they data or executables. Synchronization is kept by taking the input of the conference participant with the floor and replicating it at all other sites. In a LAN environment, it might have been easier to have taken a centralized approach and to have kept one copy of the files, running the application in one place and duplicating the output to all sites. In a geographically distributed environment, a centralized system may result in unacceptable communication delays. Centralized architectures may provide poor interactive response to the conferee with the floor when accessing an application running at a different site [24]. They may also have the drawback that they impose a heavier level of network traffic than replicated architectures because output, rather than input, must be distributed to all sites. These disadvantages are masked over LANs, due to low delays, but they are exacerbated by the large distances involved in transcontinental WANs. Because of this MMConf's strategy seems more suited to the WAN setting.

Because of MMConf's replicated architecture, applications must avoid operations that are dependent on the timing of input. To avoid nondeterminism, applications are specifically designed to run under MMConf. To allow the integration of arbitrary applications, MMConf includes the Viewshell application. Unlike other MMConf programs, it behaves like a centralized architecture and only runs at one site. It allows any program that uses simple character input/output to be run from within it, so MMConf users may import whatever such applications they want, though interactive response may suffer. Because the window management facilities are not available within the Viewshell window, graphically oriented programs cannot be used. For this reason, Lauwers *et al* lobby for centralized computer conferencing architectures, but they conclude that modern window systems make this task very difficult [25, 23].

Whether a replicated or centralized architecture is chosen, there are other difficulties of developing computer conferencing systems for operation in large diverse communication environments like the Internet. For example, the Internet may at times provide highly variable delays or routing failures that create brief service outages. Computer conferencing systems cannot function unless the underlying networks provide robust communication.

### 3.4 Video Walls

An important trial implementation of teleconferencing was the video wall experiment conducted by Xerox Corporation. Two research facilities, one in Portland, OR, the other in Palo Alto, CA, were linked via a digital

video channel and two standard half-duplex phone lines. The omnipresent connection, operating continuously 24 hours a day, encouraged unplanned interactions across the two sites. Although an additional line was installed for access to data via the Internet, no formal groupware tools were used at the time to coordinate joint work. Large format monitors were placed in common areas and limited office-to-office connectivity was provided through switching. Preliminary data indicated that 70% of all communications were of a casual, drop-in nature, with users reporting that most probably would not have occurred in the absence of the video link [16]. Roughly two-thirds of all interactions were primarily technical in nature, the remainder being social [31].

Another teleconferencing experiment, called Video Window, is being conducted by Bellcore between Morristown and Red Bank, New Jersey. Dual video channels displayed on side-by-side projection screens with careful merging provide a wide aspect ratio to view several participants. The high-quality video is coupled with quadrasonic audio. The system is expected to be up all the time, so that, like the Xerox video walls, one can just walk into the conference room to use it.

### 3.5 Cruiser

Like the video walls, Bellcore's Cruiser project [31] focuses on real-time audio and video, and specifically caters to unplanned, informal interactions. In contrast, Cruiser targets inter-office teleconferencing. Cruiser currently operates within a single research site and, since this research is not trying to solve the communication problems associated with live digital media, transmits its audio and video in analog form separately from computer data. Both video walls and Cruiser attempt to overcome the disadvantages posed by the lack of physical proximity. However, Cruiser caters not so much to continuous audio-video presence, but rather to *personal conferencing*, where individuals rendezvous as desired.

The motivations for the Cruiser project were studies that suggest physical proximity between scientific researchers leads to research collaborations [21]. The notion is that physical proximity invites frequent and spontaneous communication that in turn often initiates collaboration. The correlation is even found to extend to informal encounters, of a social nature. Furthermore, distances as little as "around the corner", "over one hallway", or "on the next floor", all within the same building, were enough to hinder potential communication between co-workers. Therefore, personal conferencing would seem to benefit not only individuals separated by large geographic distances, but also those separated by the often haphazard layout of a typical office building.

The inspiration for the project came from the George Lucas film "American Graffiti". Both focus on social interaction, although Cruiser refers to automated social browsing via a desktop computer. The implication is that personal conferencing includes not only task specific meetings, but also multimedia encounters of a purely social variety. Cruiser's interaction protocols reflect a vast array of real-world social protocols. A preference for closing one's office door or the attitude only-interrupt-me-if-it's-my-boss have corollaries in the electronic realm, as do purely technology-age phenomena such as answering machine tag.

### 3.6 The DARPA Multimedia Conferencing Project

The Multimedia Conferencing (MMC) project, a collaborative effort between ISI and BBN STC under DARPA sponsorship, has developed an experimental system for real-time, multisite conferences [6, 33]. While some conferencing projects have focused on issues for same-room conferencing, and others have concentrated on inter office conferencing over LANs, our work has instead targeted remote conferencing across transcontinental packet switched networks. By coupling real time voice and video with the MMConf computer-based shared workspace (described earlier), the system allows geographically separated individuals to collaborate. MMC is typically used for scheduled tele meetings often lasting all day. The links are not operated continuously as was done in the video wall experiments.

The underlying communication framework for all media is packet switched. The current project grew out of an initial interest during the 1970's in research on packet transmission of voice. This evolved into an interest in packet switched video as a good application for stressing the network. The system relies on an experimental suite of protocols for real time data [13, 38] and the underlying experimental Terrestrial Wideband Network (TWBnet). Packet switching technology promises to allow improvement of video quality by efficiently supporting variable-rate video coding. Its inherent multiplexing of multiple streams also allows more efficient multi-destination delivery for N-way conferencing [6]. However, most video codecs are designed for dedicated circuits, so part of the work for MMC was to adapt these codecs to packet switching.

A disadvantage of packet-switching is the potential for an increase in communication delay. In an earlier version of the system when the packet video data was sent over a satellite network, a half-second end-to-end delay could be noticed. Communication delays caused delayed reactions by users. If nothing else, participants learned to be etiquette conscious. If not, more interruptions occurred, since a pause in the middle of a sentence was difficult to distinguish from the end of it. It was especially precarious to tell a joke and have to wait a few seconds for a response. Since the network is now terrestrial, the delay has been reduced to about a hundred milliseconds, but users still sometimes notice the delay.

The system originally provided point-to-point communication. Later it was extended to support multisite conferences and now allows video from up to four sites to be simultaneously displayed in quadrants on the video monitor. At each site, voice data from remote sites is mixed for playback, allowing all sites to talk at once if they wish. Approximately half of all meetings involve more than two sites. Teleconferencing meeting room facilities currently exist near Boston, Los Angeles, San Francisco, Washington, D.C., and London, England.

A handicap of the MMC meeting room approach is that, unlike Colab, each site is only configured with one workstation to run MMConf. Ideally, all participants at all sites should be equipped with a workstation since this was the premise on which MMConf was built. Due to the bulkiness of the monitors and the limited field of view of the room cameras, this has not been practical. As a result, one participant plays scribe at each site.

The current direction of the MMC project is to provide office-to-office conferencing in addition to already-supported meeting room conferencing. Inter-office conferencing will allow individuals to be surrounded by their usual assemblage of working tools, reference materials, and other familiar resources. The MMC project is moving towards personal conferencing by porting the real-time processing components of the system onto a workstation. Although some studies [9] indicate a preference for inter-office conferencing over conference room teleconferencing, meeting room style conferencing also has its place. It has been noted that meeting rooms better accommodate groups of more than three conferees and that they typically provide higher audio quality.

### 3.7 Other Systems

Numerous other multimedia conferencing systems have been developed in the last few years. Noteworthy conferencing implementations include work at SRI [1], the Rapport system at Bell Labs [2], several MCC efforts [10], the Olivetti-sponsored research of Lantz and Lauwers [25] as well as the Pandora project [19], and Sakata's work at NEC [32].

## 4 Observations

A myriad of factors have given rise to the diversity in conferencing systems. Yet, certain themes are pervasive. The following observations have been made about the successes and failures in providing effective interfaces for multimedia conferencing.

**Familiarity.** Both Colab and Capture Lab stressed the benefits of incorporating familiar elements from conventional meeting rooms, and the inter-office emphasis of the Cruiser project promotes the usefulness of allowing conferees the resources of their natural surroundings. With familiarity come spontaneity and informality, both ingredients for making an environment conducive for collaborations. Similar arguments are made by those who favor conferencing systems that allow one's usual collection of desktop tools to be used [4, 9, 17].

**Cognitive overload.** Designers of the Capture Lab were worried about the impact of simultaneous typing, that the "cognitive load" would be too high for participants to carry on useful conversations if everyone could enter data at the same time. Consequently, they shifted away from Colab protocols to a floor policy of sequential data entry. Likewise, multiple video sources may present cognitive hazards in teleconferencing systems. To intelligently organize and even manage several real time images may overwhelm the average user. Some current commercial systems employ a separate operator to manage these details so participants are free from such worries. It remains to be seen how cognitive overload will ultimately effect the manageability, not to mention scalability of multisite teleconferencing systems.

**Simplicity.** In certain settings, simplicity of the shared applications is deemed a plus. This was certainly found to be the case at ISI, where groups of users convene in a special meeting room and then conference between meeting rooms sharing a group account. Comments about the need for simplicity were in reaction to a sophisticated

multimedia editor [14] that was the original multi user application provided with the system. Surprisingly, the criticism was aimed at its wonderfully rich functionality. Many users preferred an interface they could intuitively operate (e.g., draw with the mouse, click on the screen to change cursor position, enter text by simply typing). In response, a much simplified, rudimentary text editor was later provided for those not familiar with the fancy editor, those without time to learn it, and those simply uninterested in learning yet another new application. This sentiment was echoed by Capture Lab designers who deliberately made "the number of commands a user needed to know and the number of steps a user took to perform any type of communication task as low as possible" [26]. This was also evidenced at Colab by the distillation of one of the original applications into a simpler version with little to no use of menus [15].

**Flexibility.** Different kinds of meetings will arise. They will call for a variety of applications, different levels of sophistication among users, and a range of floor control options. Some will be informal, such as developer sessions or the gathering of an engineering task force, while others will be more formal, such as funding negotiations. As mentioned earlier, Capture Lab observed three types of face-to-face meeting styles. Similarly, Sakata studied the usage of his in house multimedia desktop computer conferencing system [32] for a year and discovered two types of meetings, one of a brainstorming nature, the other of a broadcast nature with chairperson-controlled floor changes. Recognizing the diversity of meeting situations and audiences, Colab developers had designed a range of tools from formal to informal, resulting in a total of 16-18 applications as of August 1989 [15]. Similarly, the latest version of MMConf now offers a choice of floor policy options. In short, it is necessary to offer interfaces that are appropriate for the different usages of conferencing systems.

**Necessity of real-time media.** Remote conferencing is impractical without some form of real-time voice communication. Yet does a visual channel add significant content to a teleconference? While video is not the most important information conveyed during tele meetings, it provides otherwise unavailable visual cues that help in correctly gauging reactions, in establishing a better working rapport, and in noting if conference participants are attentive, not to mention awake! Although some studies have pointed to the dubious nature of adding a visual channel, others have found there is more of a demand for video if it can be provided cheaply enough [9]. Video acts not only as a means of viewing remote co-workers, but also as a means for providing the three dimensional equivalent of a scanner or facsimile. Cameras are able to digitize any three dimensional item, as well as items that might be damaged by a pass through a scanner, such as old manuscripts and paintings [39].

**Quality of real-time media.** In the Xerox video wall experiments, "despite mediocre quality of both audio and video, users reported that the system was moderately useful for sharing culture and maintaining relationships across the two sites" [31]. However, the 56 Kbps digital video channel was considered insufficient for crucial aspects of joint work, such as detailed collaboration or delicate negotiation [21]. Audio ambiance is perhaps even more important than video quality because most information is carried in the audio channel. First reactions to the MMC system often include favorable comments on audio quality, and in particular on the ability of all sites to generate and hear audio simultaneously. By contrast, a traditional speaker phone call provides a half-duplex connection. Our personal experience has been that half duplex audio severely hampers group dialogue since it alters natural face to face group protocols and makes sizable group conferences more cumbersome. As research associates, our tolerance of reduced quality may differ from that of a business executive or a funding sponsor. Even high-fidelity systems will always be subject to criticism since teleconferencing facilities can never entirely match the quality of "being there". Because the successful existing systems span a wide range of audio and video quality, it would appear that acceptance of these systems has more to do with "the nature of the intended application than with the details of technical quality" [9].

**Communication delays.** Concern about minimizing communication delays seems to come in two flavors. concern about delays for collaborative tools to propagate updates to all sites, and about end-to-end delay of real-time media. Minimizing interaction delay is at the root of the centralized versus replicated debate for computer conferencing architectures. Delays may be noticed during information updates, as well as during floor changes. End to end delay of real time audio has the potential to effect normal conversation, as seen when MMC traffic was sent via satellite, delay is generally undetectable when under twenty milliseconds, can cause trouble when significant echo is present if between 40 80 msecs, and begins to effect normal conversation when greater than the hundred millisecond range [37].

**Technology and human factors.** Foster states "the mission of the Colab project is how computers can make groups more effective at their work" and that this requires an awareness of the interplay between technology and

group productivity [15]. Details as seemingly unimportant as inter-participant viewing and changes to room accoutrements effected acceptance of the Capture Lab system. Video monitor height and the size of video images figure into people's receptiveness to the MMC teleconferencing system. When MMC is used in fullscreen display mode, rather than quadrants, and the live video of the remote site takes up the entire video monitor, images are closer to life-size and there is more a sense of holding a face-to-face meeting. Cruiser's sensible range of social protocols (from "anyone may drop in" to "no one is to bother me") serve to offset Orwellian reactions to the presence of cameras [3, 9, 39].

## 5 New Technology

Technology advances at a rapid pace. We now expect workstations to come with a mouse. We are beginning to expect workstations to come with audio. Workstations such as the NeXT and SPARCstation already have built-in audio input and output capabilities including digitization and compression hardware and/or software. Workstations can already display video in a window with the addition of a peripheral card, such as the Parallax video card [28] and others. Within a few years, we expect workstations to come with integrated support for motion video including built-in cameras and video bandwidth compression hardware.

The inclusion of video capabilities within the workstation architecture will allow video to be treated like other more traditional forms of data, able to be stored on conventional file systems, edited, and shipped over the network [19, 39]. Widespread availability of and familiarity with integrated workstation video will also lead to a surge in multimedia applications since developers will be able to depend on video capabilities being there.

Video produces a staggering amount of raw data for each frame, and many frames per second, producing a data stream of approximately 100 Mbps. Compression is essential not only to lessen network loads, but is also critical for storage. Researchers continue to improve bandwidth compression algorithms [30]. Motion video can now be transmitted at 56 Kbps. As compression standards are established, these functions will be implemented in VLSI and incorporated directly into workstations.

Workstations already incorporate high speed packet-network interfaces, such as Ethernet, and the speed of these networks will increase, for example to 100 Mbps with FDDI and other high-speed LANs [34]. There are already many 1.5 Mbps networks to connect LANs, some networks, such as the NSFNET backbone, are being upgraded to 45 Mbps, and the recent government initiative to establish the gigabit NREN will bring high-speed networking to a much larger sector of the population. This network infrastructure in combination with video capabilities in the workstation will enable widespread multimedia conferencing, to support projects such as the Collaboratory.

Support for real time data over these networks will require real-time communication protocols similar to those the MMC project has developed [6, 38]. These protocols provide for low delay transmission and multicast delivery for multisite conferences. Multicast delivery is also valuable for non-real-time protocols [8] used in groupware applications.

## 6 Conclusions

Technology alone cannot transform electronic conferencing into an accepted or widespread form of communication. Each conferencing project surveyed had its own discoveries of seemingly unimportant human factors that flew in the face of technology—small modifications of an either psychological or sociological import that seem to propel the cause forward. A modicum of success is implied by the continued use of these systems. If nothing else, their usage reflects a good match between the capabilities of the systems and the tolerance, expectations, and needs of their user communities. An integral part of the coming-of-age process will be the continued attention to issues beyond the scope of technology itself.

The idea that collaborative technology is an activity in search of a need should be laid to rest. Its niche is among individuals who spend most of their time in group endeavors, who use computers to do their work, and whose potential for collaborations has been impaired by lack of geographic proximity. It seems especially well suited for the kinds of scientific collaborations envisioned for the Collaboratory. The trend among the multimedia conferencing system: discussed is to draw on real world interaction protocols, but not to enforce a strict electronic replica of face to face conferences. Instead, multimedia conferencing is promoted as a supplement to face-to-face collaboration.

Each of the systems has focused on different issues, which in turn has given rise to a set of at once varied and recurring findings. Colab research in user interfaces for simultaneously-accessible shared workspaces has been instrumental in establishing electronic metaphors for group interactions. Capture Lab stresses the influence of human factors on the acceptance of conferencing systems. Video walls are found to be effective as a means for co-presence, while Cruiser takes this one step further with personal conferencing -- where unplanned, informal interactions and social browsing are considered virtues along the road to research collaborations. MMConf, and systems like it, introduce the idea of groupware in remote settings and the architectural tradeoffs for supporting joint work over large geographical distances. Finally, MMC integrates real-time media with remote conferencing over a wide area network.

Anecdotal rules of trade have been distilled from the recurring themes observed. Simple considerations, such as accommodating a variety of conferencing scenarios, guarding against cognitive overload, and catering to a sense of familiarity, have repeatedly been cited as guidelines used by system builders. The necessity and quality of real-time media also figure into a system's effectiveness, as do the simplicity of groupware interfaces and the impact of communication delays.

In conclusion, multimedia conferencing has not yet come of age. The criticisms are being addressed, and the underlying technology is nearly ready. Future technological developments will provide the higher speed networks and more sophisticated workstations needed to make multimedia conferencing feasible. A wider availability of such systems will result in a growing corps of users less intimidated by the technology. Complaints about response times, network bandwidth, inflexible shared window systems, and hardware bulkiness, all have solutions on the horizon. Prohibitive costs and the undue complexity of current day systems should be eliminated by next-generation workstations. In turn, this would leave more time to devote to other pressing questions. When are computers and live media suited to group conferencing? What group tasks are appropriate for electronic mediation? How can we create systems that seamlessly transition between one's personal work environment and conferencing workspaces?

### Acknowledgement

This research was sponsored in part by the Defense Advanced Research Projects Agency under Contract Number DABT63-91-C-0001. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency of the United States Government.

### References

- [1] Aguilar, L., Garcia Luna Aceves, J., Moran, D., Craighill, E., Grungardt, R., "Architecture for a Multimedia Teleconferencing System", *ACM Proceedings SIGCOMM '86 Symposium on Communications Architectures and Protocols*, Stowe, VT, pp.126-136 (Aug 1986).
- [2] Ahuja, S.R., Ensor, J.R., Horn, D.N., "The Rapport Multimedia Conferencing System", *Proceedings of the Conference on Office Information Systems*, ACM-SIGOIS, pp.1-8 (Mar 1988).
- [3] Bell Systems Research Labs, "The Picturephone System", *Bell Systems Technical Journal* (Feb 1971).
- [4] Bikson, T., Presentation at IFIP Working Group 8.4, session on "The Present and Future of Groupware", Palo Alto, CA (Aug 1989).
- [5] Bush, V., "As We May Think", *Atlantic Monthly*, Vol.176, No.1, pp.101-108 (June 1945).
- [6] Casner, S., Seo, K., Edmond, W., Topolcic, C., "N Way Conferencing with Packet Video", *Proceedings 3rd International Workshop on Packet Video*, VISICOM '90, Morristown, NJ (Mar 1990).
- [7] Crowley, T., Milazzo, P., Baker, E., Forsdick, H., Tomlinson, R., "MMConf. An Infrastructure for Building Shared Multimedia Applications", *Proceedings CSCW '90*, Los Angeles, CA (Oct 1990).
- [8] Deering, S., "Host Extensions for IP Multicasting", RFC 1054, Stanford University, Stanford, CA (May 1988).
- [9] Egidio, C., "Video Conferencing as a Technology to Support Group Work. A Review of its Failures", *Proceedings CSCW '88*, Portland, OR, pp.13-24 (Sept 1988).
- [10] Ellis, C., Gibbs, S., Rein, G., "The Groupware Project. An Overview", Technical Report STP-033-88, Software Technology Program, MCC, Austin, TX (Jan 1988).
- [11] Engelbart, D.C., "Collaboration Support Provisions in AUGMENT", OAC 1984, *Proceedings of 1984 AFIPS Office Automation Conference*, pp.51-58 (Feb 1984).
- [12] Engelbart, D.C., "NLS Teleconferencing Features. The Journal, and Shared Screen Telephoning", *Fall COMPCON 75 Digest of Papers*, pp.173-177 (Sept 1975).
- [13] Forgie, J.W., "IEN 119. S1 A Proposed Internet Stream Protocol" (unpublished memorandum), MIT Lincoln Laboratory, Cambridge, MA (Sept 1979).
- [14] Forsdick, H.C., "Explorations into Real time Multimedia Conferencing", *Proceedings 2nd International Symposium on Computer Message Systems*, pp.299-315 (Sept 1985).

- [15] Foster, G., "Tools for Real-time Collaborative Work in the Colab", position paper and demonstration at the Groupware Technology Workshop, IFIP Working Group 8.4, Palo Alto, CA (Aug 1989).
- [16] Goodman, G., Abel, M., "Communication and Collaboration. Facilitating Cooperative Work through Communication", *Office. Technology and People*, Vol.3, No.2, pp.129-146 (1987).
- [17] Grudin, J., "Why CSCW Applications Fail. Problems in the Design and Evaluation of Organizational Interfaces", *Proceedings CSCW '90*, Portland, OR, pp.85-93 (Sept 1988).
- [18] Gust, P., "Multi-user Interfaces for Extended Group Collaboration", Groupware Technology Workshop, IFIP 8.4 Working Group, Palo Alto, CA (Aug 1989).
- [19] Hopper, A., "Pandora An Experimental System for Multimedia Applications", *ACM Operating Systems Review* (Apr 1990).
- [20] Ives, H.E., "Two-way Television", *Bell Labs Record*, Vol.8 (1930).
- [21] Kraut, R., Egido, C., Galegher, J., "Patterns of Contact and Communication in Scientific Research Collaboration", *Proceedings CSCW '88*, Portland, OR, pp.1-12 (Sept 1988).
- [22] Lai, K.Y., Malone, T.W., "A 'Spreadsheet' for Cooperative Work", *Proceedings of CSCW '88*, Portland, OR, pp.115-124 (Sept 1988).
- [23] Lantz, K., Lauwers, C., Arons, B., Binding, C., Chen, P., Donahue, J., Joseph, T., Koo, R., Romanow, A., Schmandt, C., Yamamoto, W., "Collaboration Technology Research at Olivetti Research Center", Groupware Technology Workshop, IFIP Working Group 8.4, Palo Alto, CA (Aug 1989).
- [24] Lantz, K., "An Experiment in integrated Multimedia Conferencing", in *Computer Supported Cooperative Work. A Book of Readings* (I. Greif, Ed.), Morgan Kaufmann Publishers, Inc., San Mateo, CA (1988).
- [25] Lauwers, J.C., Joseph, T.A., Lantz, K.A., Romanow, A.L., "Replicated Architectures for Shared Window Systems. A Critique" *Proceedings of the ACM Conference on Office Information Systems* (Mar 1990).
- [26] Mantei, M., "Capturing the Capture Lab Concepts. A Case Study in the Design of Computer Supported Meeting Environment", *Proceedings CSCW '88*, Portland, OR, pp.257-270 (Sept 1988).
- [27] Panko, R.R., "Office Work", *Technology People*, Vol.2, pp.205-353 (1964).
- [28] Parallax Graphics, *The Parallax 1280 series Videographic Processor*, Parallax Graphics Technical Report (1987).
- [29] Postel, J., Finn, G., Katz, A., Reynolds, J., "The ISI Experimental Multimedia Mail System", ISI Research Report ISI/RR-86-173, USC/Information Sciences Institute, Marina del Rey, CA (Sept 1986).
- [30] Packet Video Workshop, Session C. Coding, *Proceedings 3rd International Packet Video Workshop*, Morristown, NJ (Mar 1990).
- [31] Root, R.W., "Design of a Multi-Media Vehicle for Social Browsing", *Proceedings CSCW '88*, Portland, OR, pp.25-38 (Sept 1988).
- [32] Sakata, S., "Development and Evaluation of an In-House Multimedia Desktop Conference System", *IEEE Journal on Selected Areas in Communications*, Vol.8, No.3, pp.340-347 (Apr 1990).
- [33] Schooler, E.M., Casner, S.L., "A Packet-switched Multimedia Conferencing System", *ACM SIGOIS Bulletin*, pp.12-22 (Jan 1989).
- [34] Schroeder, M.D., Birrell, A.D., Burrows, M., Murray, H., Needham, R.M., Rodeheffer, T.L., Satterthwaite, E.H., Thacker, C.P., "Autonet. A High-speed, Self-configuring Local Area Network using Point-to-Point Links", Research Report 59, DEC SRC, Palo Alto, CA (Apr 1990).
- [35] Snyder, F.W., "Travel Patterns. Implication for New Communication Facilities", Bell Laboratories Memorandum (1971).
- [36] Stefik, M., Foster, G., Bobrow, D.G., Kahn, K., Lanning, S., Suchman, L., "Beyond the Chalkboard. Computer Support for Collaboration and Problem Solving in Meetings", *Communications of the ACM*, Vol.30, No.1, pp.32-47 (Jan 1987).
- [37] Swinehart, D.C., Stewart, L.C., Ornstein, S.M., "Adding Voice to an Office Computer Network", *Proceedings IEEE GlobeCom '83 Conference* (Nov 1983).
- [38] Topolcic, C., Park, P., Draft of "Proposed Changes to the Experimental Internet Stream Protocol", BBN Laboratories, Cambridge, MA (Apr 1987).
- [39] Wilson, I.D., Milway, D., Hopper, A., "Experiments in Digital Video for Workstations", Technical Report ORL-89-2, Olivetti Research Limited, Cambridge, England (Mar 1989).